RESEARCH NOTE NUMBER 418

SUMMARY OF
SPECTRAL SCANNING THEORY OF COLOR VISION*

By

George Biernson

UNPUBLISHED PRELIMINARY DATA

9 SEPTEMBER 1963

APPLIED RESEARCH LABORATORY
SYLVANIA ELECTRONIC SYSTEMS
A Division of Sylvania Electric Products Inc.
40 SYLVAN ROAD, WALTHAM 54, MASSACHUSETTS

*The work reported herein was sponsored under the following contracts: AF-33(657)-8681 with the Air Force Systems Command, United States Air Force, and NASW-441 with the National Aeronautics and Space Administration.

SPECTRAL SCANNING THEORY OF COLOR VISION

By

George Biernson

It has generally been assumed that color vision is a three dimensional process. The eye presumably derives from the incoming spectra three independent signals, and somehow inter-relates these three signals to produce the color sensation. The retina is assumed to have three types of photosensitive elements (either in different types of cones or in different parts of a single cone) which have different spectral absorption responses, in accordance with the theory proposed by Thomas Young[1] in 1801.

Numberless theories of color vision have been proposed to explain how this principle is implemented in detail. The most popular today are the so-called "stage" theories, which postulate that the three signals from the photosensitive elements are combined together in some manner to form at a later stage of the visual process three opponent-process signals. The three opponent-process signals represent the sensations, white-black, yellow-blue, and red-green. One of the colors of each pair is evoked when the signal is positive and the

other color when the signal is negative. That the eye senses colors in terms of these opponent-process signals was first proposed by Ewald Hering[1] in 1878.

It is interesting to note that the NTSC color television transmitting system now used in commercial broadcasting operates essentially in accordance with the stage theories of color vision. By means of filters the image is separated into red, green, and blue pictures. These are focused onto three separate television camera tubes, which act as the three photosensitive elements of the Young theory. For each point on the image, separate red, green, and blue signals are produced. These three signals are combined together in a mixing network to from opponent-process signals approximately corresponding to white-black, blue-yellow, and green-red. Black-and-white television receivers accept only the white-black or luminosity signal, whereas color television receivers also detect the yellow-blue and red-green chromaticity signals.

Almost all theories of color vision have assumed that the receptor process is three dimensional. A few color theorists (such as Hartridge[1]) have assumed more than three types of photosensitive elements, which is equivalent to assuming a dimensionality greater than three; and a few (such as Houstoun[2]) have postulated that the visual process deals with

continuous data. However these theories have not been taken very seriously, primarily because (1) there is great difficulty in reconciling them with constancy of three-dimensional metameric color matches, (2) no strong case has been presented to show a need for more than three dimensions in color vision and (3) the physiological models proposed appeared to be less acceptable than the simpler concept of Thomas Young.

This paper proposes a new reason for rejecting the three-dimensional concept of color vision. There is serious question whether any three-dimensional spectral discrimination system could achieve the extreme accuracy and wide adaptation range of human vision. Judd* has reported that the eye can discriminate among ten million shades of object color. The retina can adapt to cover a range of light intensities of one billion to one, and over a significant part of this range has color vision. The retina automatically adapts to large changes of intensity and spectrum of the illuminant so accurately that we hardly know the illuminant has changed, and we think of object color as an inherent characteristic of the object, independent of the illuminant. It does not seem possible that the electronic system employed in color television cameras (which uses the principle on which color vision is assumed to be based) could ever be extended to even approach the performance

*Reference 3, page 171.

of human vision. The neurological data processing of the color signal in the brain is not a factor in this question, because the brain cannot perform accurate spectral discrimination unless the retina itself is capable of providing data of the required accuracy.

Thus, the question to ask is not "How does the eye see color?", but rather, "How could the eye see color with the accuracy of spectral discrimination that we know it has?"

A reasonable explanation for the spectral discrimination accuracy of color vision presents itself if we assume that each cone of the eye acts as a spectrum analyzer and performs visual adaptation in terms of spectral information. As will be shown, such a system could achieve by means of simple feedback processes very high accuracy of spectral discrimination and the ability to compensate for wide variations in the illuminant. The spectrum analyzer could be very crude, and have foldover of spectral information and distortion of the wavelength scale, because it is the feedback processes that provide the accuracy of spectral discrimination.

To act as a spectrum analyzer, each cone must have a prismatic effect that tends to concentrate different wavelengths in different regions of the cone, and a scanning process that scans across the cone to read off the spectral

information. The scanning process could be performed by chemical or electrical effects on a molecular level and so might not be detectable by microscopic examination of the cone. The prismatic effect could be produced by waveguide made patterns in the cylindrical photosensitive region of the cone. We know that waveguide mode effects produce different spatial distributions of energy as a function of radius for different wavelengths, and these distributions vary smoothly as the wavelength is varied. The waveguide mode patterns qualitatively satisfy the requirements of the prism for our spectrum analyzer, but considerable research remains to be done to show whether there is quantitative agreement.

To see how the eye might adapt in terms of spectra, let us consider first the photodetection process. When a photon of light is absorbed by a photopigment molecule in a visual receptor, it triggers a chemical reaction, which eventually results in what is called bleaching of the molecule. The bleached molecules revert back to the unbleached state in a random manner, such that the average rate of regeneration of unbleached photopigment molecules is proportional to the concentration of bleached molecules.

Let us assume that a visual receptor acts as a balanced electrical bridge, and that a unit of current flows in an arm of the bridge every time a photopigment molecule is bleached or

regenerated. The currents due to bleaching would flow in one pair of arms and the currents due to regeneration would flow in the other pair. The effect would be for the regeneration currents to act as a self-adjusting balancing mechanism, which adapts the receptor in an ideal manner to any level of illumination. The signals from the detection process would be

$$S = \frac{B - R}{B + R} = \frac{L - L_{av}}{L + L_{av}} \qquad (1)$$

where B is the rate of photopigment molecule bleaching, R is the rate of photopigment molecule regneration, L is the luminosity that the receptor is now sensing and L_{av} is the time average luminosity to which the receptor has become adapted. The bleaching rate B is proportional to the instantaneous luminosity L and the regeneration rate R is proportional to the average luminosity L_{av}.

Equation (1) was empirically derived by Adams and Cobb[4] in 1922 (in a slightly different form), and has been shown by Judd* to give a good description of visual adaptation. Yilmaz[5] concluded that visual adaptation must follow a mathematical law equivalent to the Lorentz transformation of Special Relativity in order for the eye to compensate accurately for changes

*Reference 3, page 226.

in illuminant, and Equation (1) follows this mathematical requirement.

Let us assume that the balanced electrical bridge circuit is a localized effect within the cone. In a small region of the cone, a localized regeneration current balances the effect of a localized bleaching current. Since the energy for different wavelengths is concentrated at different regions of the cone, this allows the cone to adapt in a different manner for different wavelengths. The effect is that the cone achieves essentially ideal adaptation in terms of the spectrum of the light.

When a receptor is excited by a particular spectrum, waveguide mode effects transform that spectrum into a spatial energy distribution in the receptor. This produces a spatial distribution of bleaching currents proportional to the energy distribution. If the excitation spectrum is kept constant, a compensating spatial distribution of regeneration currents builds up to balance the effect of the bleaching current distribution.

Normally the eye is continually moving, and so each cone becomes adapted to the average spectrum in the field of view. In a normal scene this average spectrum can be considered as

being approximately proportional to the spectrum of the illuminant. Thus a regeneration current distribution would be produced in each cone that approximately compensates for the spectrum of the illuminant. The net signal produced in the detection process would therefore correspond approximately to the reflectivity spectrum of the object being sensed by the cone. This would allow the cone to compensate very accurately for the spectrum of the illuminant and provide a signal defining the reflectivity spectral characteristics of the object itself.

It is postulated that a chemical or electrical process scans across the receptor, detecting in each local region the output from its balanced bridge circuit. This generates in each cone a modulated waveform that approximately defines the reflectivity spectrum of the object, where the time variable defines wavelength. There is reason to believe that the scanning process scans at a rate of about 20 cps.

It is postulated that the modulated waveform from each cone is compared with a weighted average of the modulated waveforms from the other cones of the retina. The waveform from the cone would be modified to accentuate differences between it and the weighted average waveform, which is equivalent to accentuating differences between the reflectivity spectrum of a point and the average reflectivity spectrum over the field

of view. Thus spectral data processing would provide what is called contrast enhancement, and would allow the retina to detect very slight differences in object reflectivity spectra.

The modified waveform from each cone is then demodulated to form the three-dimensional opponent-process color signals. The average value gives the white-black or luminosity information and the modulation components give the chromaticity information. The yellow-blue signal is obtained by essentially subtracting the energy in the time interval of the modulated waveform corresponding to short wavelengths from that corresponding to long wavelengths. The red-green signal is obtained by essentially subtracting the energy corresponding to the center of the visual wavelength range from the energy corresponding to the ends of the visual range.

The cone delivers to different bipolar cells three separate voltages, which represent the white-black, yellow-blue, and red-green opponent-process signals. These voltages are continuous and have positive and negative values. (The opponent process voltages could be superimposed upon a bias level in the receptor so that the net voltage has only one polarity.) A single bipolar cell combines together one type of signal (i.e., a white-black, a yellow-blue, or a red-green) from a number of

different receptors, and feeds the combined signal to ganglion cells. A ganglion cell converts the continuous voltage into nerve impulses, which is the signal form needed to allow the information to be "telemetered" to the brain along the optic nerve.

Our classical theories of color vision place strong emphasis on the phenomena of three-dimensional color matches. Two lights can have entirely different spectra and yet evoke the same color sensation in the eye. Such lights are said to be metamers of one another. It is found that if two spectra match they will continue to match approximately regardless of the color of the background light that surrounds them, even though the colors evoked by the metamer spectra change as the background color is changed. This phenomenon is called constancy of metameric match.

The concept that color vision is a three-dimensional process is based primarily on the following color matching experiment. A subject observes two adjacent patches of light formed by combinations of a sample spectra and a number of standard lights called primaries, and he adjusts the intensities of the primaries until the colors of the two patches look the same. In this experiment he may add one of the primaries to the sample

to form one patch and match this modified sample spectrum to admixtures of the other primaries. It is found that if the primaries are metamerically independent (i.e., if they cannot be mixed to match themselves), only three primaries are needed to match all possible spectra.

How can the Spectral Scanning theory explain the three-dimensional color matching experiment? All that is required is that the spectral data processing performed in the cone must be approximately linear and only three-dimensional information concerning the spectrum is sent to the brain. Such a system would approximately satisfy constancy of metameric matches and related experiments, which is all that is required.

The key issue here is the question of accuracy. A metameric matching experiment may be accurate relative to many psychophysical measurements, but it is extremely crude in comparison to the ability of the eye to discriminate among objects with similar reflectance spectra. Despite the strong emphasis placed on metameric matches in classical color vision theories, little effort has been spent to study the accuracy with which metameric matches hold.

The metameric match phenomenon really defines a limitation of the color vision mechanism rather than its foundation. To perform its task of protecting the organism, the eye must be

able to discriminate accurately among small differences in object reflectance spectra. The mere fact that two incoming light spectra may look alike even though they are quite different is merely a biproduct of the means used for discrimination. Thus the eye must have very high accuracy in observing differences in spectra that are nearly alike, but need not have accuracy in performing metameric matches.

Classical color vision theories have proposed techniques that could accurately satisfy constancy of metameric match (which is a side issue in the visual process) but have essentially ignored the problem of accuracy in discriminating among objects with nearly the same reflectance spectra (which is the purpose for which the evolutionary forces of nature really designed vision). The theory now being proposed approaches the problem from the opposite direction. It is first necessary to find a system that can achieve very high accuracy of object spectral discrimination and then determine whether it will also satisfy the low accuracy requirement of constancy of metameric match.

This theory has been found to be in good agreement with experimental color vision data. Psychophysical color vision experiments generally fit naturally into an opponent-process color vision theory. Color-blind deficiencies are readily

attributable to defects in the waveguide mode patterns in the cones. Although spectrophotometer measurements of individual receptors and of the intact retina (such as those performed by Rushton) have been interpreted as proving the existence of more than one photopigment in the cones, these measurements are also consistent with the concept of a single type of cone with a single photopigment. The measurements can be explained satisfactorally in terms of the effects of different waveguide modes in the receptor.

It is interesting to note the strong similarity between this theory and the following postulate proposed by Edwin Land[6] a few years ago:

> "What happens in the real world, when the eyes receive a continuous band of wavelengths? We are speculating about the possibility that these wavelengths register on the retina as a large number of individual color-separation "photographs," far more than the three that Maxwell thought necessary and far more than the two that we have shown can do so well."

A detailed development of this theory is presented in reference 7.

REFERENCES

1. R.C. Teevan and R.C. Birney, Color Vision, (D. Van Nostrand Co., Inc., Princeton, New Jersey, 1961).
2. R.A. Houstoun, Vision and Color Vision, (London: Longmans Green and Co., 1932).
3. D.B. Judd, Color in Business, Science and Industry, (New York: John Wiley and Sons, Inc., 1952).
4. E.Q. Adams and P.W. Cobb, J. Exp. Psych., 5, 39 (1922).
5. H. Yilmaz, Bull. Math., Biophysics, 24 (1962).
6. E.H. Land, Sci. Amer., 200 (5), 84-99 (1959).
7. G.A. Biernson, Research Report 345, Sylvania Electronic Systems, Applied Research Laboratory, Waltham, Massachusetts, June, 1963.